

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2003, 10:19:03 ; Search time 658.074 Seconds
(without alignments)
11025.472 Million cell updates/sec

Title: US-09-922-895-4

Perfect score: 448
Sequence: 1 TAGTCAGATGCAGAAAT.....CCCTGCTTAATGAAAGCTT 448

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	387.4	86.5	747	12	BG220110 RST39882
2	76.4	17.1	655	10	BB223728 BB223728
3	65.6	14.6	274	12	BF553209 UI-R-C2-n
4	49.4	11.0	320	10	BB225464 BB225464
5	47.4	10.6	1203	17	CNS015MU
c 6	47.2	10.5	1027	17	CNS02750

c 7	46.4	10.4	1101	17	CNS016LI	AL106896 Drosophila
8	46.2	10.3	914	17	CNS002JY	AL097768 Drosophila
9	46	10.3	1101	17	CNS0172Q	AL108704 Drosophila
10	45.8	10.2	975	17	CNS0102Y	AL098452 Drosophila
11	45.2	10.1	583	12	BG603853	BG603853 EST502943
12	44.6	10.0	874	17	A2541515	A2541515 ENTF060TF
13	44.6	10.0	1027	17	CNS02750	AL212733 Tetradon
14	44.4	9.9	1205	17	CNS0165A	AL106312 Drosophila
c 15	44.2	9.9	829	17	BH501921	BH501921 BOHJ748TR
c 16	44.2	9.9	987	17	CNS014PO	AL104456 Drosophila
c 17	44.2	9.9	1101	17	CNS0106X	AL098595 Drosophila
c 18	43.6	9.7	699	17	AQ323327	ALQ323327 Pct11-10
c 19	43.4	9.7	772	17	CNS03AEN	ALQ35112 Tetradon
c 20	43.4	9.7	1086	17	CNS00YXK	AL096962 Drosophila
c 21	43.2	9.6	452	13	BM273820	BM273820 PfEST0a6
c 22	43.2	9.6	1201	17	CNS016FT	AL106691 Drosophila
c 23	43	9.6	546	10	BB224066	BB224066 Drosophila
c 24	43	9.6	1101	17	CNS000D1	AL065414 Drosophila
c 25	43	9.6	1101	17	CNS00E07	AL069440 Drosophila
c 26	42.8	9.6	540	13	BJ077706	BJ077706 Drosophila
c 27	42.8	9.6	1101	17	CNS00LTX2	AL078714 Drosophila
c 28	42.6	9.5	964	17	CNS07E8R	AL441457 T7 end of
c 29	42.4	9.5	471	13	BJ413351	BJ413351 BJ413351
c 30	42.4	9.5	483	13	BJ412970	BJ412970 BJ412970
c 31	42.4	9.5	625	13	BJ417045	BJ417045 BJ417045
c 32	42.4	9.5	634	13	BJ416550	BJ416550 BJ416550
c 33	42.4	9.5	669	13	BJ413381	BJ413381 BJ413381
c 34	42.2	9.4	567	13	BJ410890	BJ410890 Drosophila
c 35	42.2	9.4	1101	17	CNS00L4E	AL067752 Drosophila
c 36	42	9.4	369	10	AW411933	AW411933 uo2908.x
c 37	42	9.4	458	17	AQ408877	AQ408877 HS_5067_A
c 38	42	9.4	499	13	BJ324359	BJ324359 BJ324359
c 39	42	9.4	556	13	BJ388900	BJ388900 BJ388900
c 40	42	9.4	656	13	BJ387997	BJ387997 BJ387997
c 41	42	9.4	956	17	CNS00KX2	AL078376 Drosophila
c 42	41.8	9.3	484	17	AQ866715	AQ866715 nbeu0028L
c 43	41.8	9.3	880	17	BH152122	BH152122 ENTQ005TF
c 44	41.8	9.3	948	17	CNS0159T	AL105179 Drosophila
c 45	41.8	9.3	951	17	AZ681741	AZ681741 ENTIP07TF

ALIGNMENTS

RESULT 1	BG220110/c	747 bp	mRNA	linear	EST 21-APR-2001
LOCUS	BG220110				
DEFINITION	RST39882	Atherys RAGE Library	Homo sapiens	CDNA, mRNA sequence.	
ACCESSION	BG220110				
VERSION	BG220110.1	GI:13746131			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	1 (bases 1 to 747) Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Dener,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,J., Denzig,J., and Ducar,M. Creation of genome-wide protein expression libraries using random activation of gene expression Nat. Biotechnol. 19 (5), 440-445 (2001) 21227151				
TITLE	JOURNAL				
MEDLINE	COMMENT				
CONTACT:	Scott J. Cain				
	Athersys, Inc.				
	3201 Carnegie Ave, Cleveland, OH 44115, USA				
	Tel: 216 431 9900				
	Fax: 216 361 9596				
	Email: scain@atersys.com				
	High quality sequence stop: 523.				
FEATURES	Location/Qualifiers				


```

OY 61 AAGCTTCACACCTCAGCTCTTAAACAGTCTCTCAAACTTCCAGTGCACACTGAAGCTC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 234 AAGTCAGTCCCAATGACCTCTAGCTCAGTCACTTGTCTCATCTGCATGCAATGCTGATGCTC 293
OY 121 TTGAAGACACTCAATATATACACAGCAGTACAGTATGATGATGATGATGATGATGAT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 294 TCAAGAGACCTGTAAACACATTAATACGAGAAATATCTCATCTACCCCAAGATCAT 353
OY 181 ACCACAGG-CCAGGGGCTGGGAGGCTACTCATCATCAACCTTAA-----AAGA 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 AGTAGAGAGAGCTGTGCTGGGAGGCTGAGCAATTTTCAATCAACGAGACACTGAAGA 413
OY 231 GAGCTTGTCTCTCTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 AAGACAGAGCTTGTGCTTATTTTATTCGTCGAAGAGTTATACATGTTCTGATTCACCT 473
OY 291 AGTACTATATCCCTCTACAAAAGTAAACTTTTATTTTATTTTATTTTATTTTATTTT 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 474 AATGTAAATATATCATTCACAGTAGAGAGTTTAAATAATTTAACTACTTTTTC 533
OY 351 CAGCTATGATATTAATAATTAATTTTACACATAATTAATTAATTAATTTTATTTT 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 534 CAGCTTTTACATCATGTAATGTT-----TCCCAAAATGCAATTTTATTTTATTTCC 584
OY 411 TAATGTCCCTAGTCTTCTTCCCTGCTTAATGAAGAAGCTT 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 585 TAACGTAAGTACTTCTTTTCTTTTGTGATGATGCTT 622

RESULT 3
LOCUS BF553209 274 bp mRNA linear EST 12-DEC-2000
DEFINITION UI-R-C2-nk-a-12-0-UI.r1 UI-R-C2 Rattus norvegicus cDNA clone
VERSION BF553209.1 GI:11662939
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 274)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL MEDLINE
COMMENT 97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@iuii.uow.edu
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LMU (infelimage.lmu.gov). IMAGE ID-1786569
Seq primer: M13 Forward.
Location/Qualifiers
FEATURES
source
1..274
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2-nk-a-12-0-UI"
/clone_lib="UI-R-C2"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73b-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The UI-R-C2
library is a subtracted library derived from the UI-R-C1

```

```

library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which allows
identification of the library of origin of a clone within
the mixture. The subtracted library (UI-R-C2) was
constructed as follows: PCR amplified cDNA inserts from
UI-R-C1 clones from which 3 ESTs had been derived was
used as a driver in a hybridization with the UI-R-C1
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was
purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-C2 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996)."

BASE COUNT 91 a 46 c 41 g 96 t
ORIGIN
Query Match 14.6%; Score 65.6; DB 12; Length 274;
Best Local Similarity 66.3%; Pred. No. 4e-06;
Matches 110; Conservative 0; Mismatches 54; Indels 2; Gaps 1;

OY 128 CACTGAATATATACACAGCAGTACAGTATGATGATGATGATGATGATGATGATGAT 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 CACGAGAGATACATTAATATGAAATATATCTTATGATGATGATGATGATGATGATG 60
OY 188 GCCAGGGCTGGGAGGCTGATCATCAACCTTAAAGAGAGAGCTTGTCTCTCT 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 GCCATGTGCTGGCCAGTATATATCATGACATGAAGAGAGAGCTTTTCTTATCT 120
OY 248 CTAAATGATTAACCTATTTTAAATGACCTGAATGTTAGTACT 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 CTAAA--GAGTATCATATGCTTTGATTCATTAATGTTAAATAT 164

RESULT 4
LOCUS BB225464 320 bp mRNA linear EST 01-JUL-2000
DEFINITION BB225464 RIKEN full-length enriched, adult male aorta and vein Mus
VERSION BB225464.1 GI:8894075
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 320)
Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,
P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Iizawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,
Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya
T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamazaki,I.,
Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
M., Kuramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Kono,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222

```



```

/organism="Drosophila melanogaster"
/db.xref="taxon:7227"
/clone="BACN02L18"
/clone_1lb="DrosBAC"
/plasmid="pBelOBAC11"
/!note="end : Sp6"

BASE COUNT      282 a      51 c      36 g      138 t      407 others
ORIGIN

Query Match      10.3%; Score 46.2; DB 17; Length 914;
Best Local Similarity 26.6%; Pred. No. 0.36;
Matches 91; Conservative 53; Mismatches 198; Indels 0; Gaps 0;

OY 103 AGTGCACACGACGACCTCTTGACGACACTGGAATGTACACACAGCAGTAGCAGTGAATGC 162
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 443 ANNNNAANAANNNNAANNNNAANNNNAANAANAANNNNAANNNNAANNNNAANNNNNNNN 502

OY 163 ATGTACCTTAAGGTGATTACACAGCGCGGCGTGGCAGCGTGTCTGATCATCAACCT 222
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 503 AANNNAANAANNAANAANAANAANAANAANAANNAANNAANAANAABAANAANNNNNNNAN 562

OY 223 AAAAAAGCAGAGCTTGTGCTCTCTCTCTAAATAGATCTACATTTTAATGACCGTGA 282
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 563 ANANNAANAANNNNNNNNNNNNNNNNAAMAHAHTTAATAAATATTTTAAATATTTTAA 622

OY 283 TGTGTGATAGTTACTATATGCGGCTACAAAAGGTAACATTTTATATTTTATATCATTA 342
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 623 TTTTATTTTATTTTATTTTAAATATATATATTTTAAATTTTAAATTAATTAHAHAAMAAT 682

OY 343 ACTTCAGCAGGATGTGATTAATTAATAAACAATTTTCACACAATACATACATAGTACTATT 402
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 YHAHTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 742

OY 403 TTATTTTCTAATGTGCTAGTTCTTTCCCTGCTTAATGAATA 444
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 743 TWAATTTTWAHAHTATTTWAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 784

RESULT 9
LOCUS      CNS01720
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
            BACN37P08 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
VERSION    AL108704
KEYWORDS   AL108704.1 GI:5629008
SOURCE     GSS.
ORGANISM   Drosophila melanogaster.
            Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
            1 (bases 1 to 1101)
REFERENCE   Genoscope.
            Direct Submisson
            Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 J1006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome project (EDGP) -
            http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBelOBAC11.

FEATURES
    Source
        1..1101
            location/Qualifiers
            /organism="Drosophila melanogaster"
            /db_xref="taxon:7227"
            /clone="BACN37P08"
            /clone_1lb="DrosBAC"
            /plasmid="pBelOBAC11"

```

[illegible]

Db 697 AMTTTWTGTTTTTTTTTMMATDAMCAAMCMHACTTAACAYCAWCAHAMMTTAMNMCMCH 756

314 AGCTAAACCTTTTATATTTTATATTAACCTACAGCCAGCTATGATATAATTAATAACA 373

Db 757 ACCTTAAATWTMTMAATWTMTTAAACGCTTTTTTTTTTATTCCTATATWTWDAAATTAAMACA 816

374 TTTTCACAAATACAAATGAATTAATTTTATTTTCTATATGCTAGTCTTTC 429

Db 817 TATTTATTTTATATATATATATTTTAAAKAKASAMATMAATTTTWTMTTWTTW 872

RESULT 11

LOCUS BG603853

DEFINITION EST502943 Plasmodium yoelii sporozoite cDNA Plasmodium yoelii cDNA

ACCESSION BG603853

VERSION BG603853.1 GI:15153867

KEYWORDS EST.

SOURCE Plasmodium yoelii.

ORGANISM Plasmodium yoelii.

REFERENCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. 1 (bases 1 to 583)

AUTHORS Kappe,S.H.I., Gardner,M.J., Brown,S.M., Ross,J., Matuschewski,K., Ribeiro,J.M., Adams,J.H., Quackenhush,J., Cho,J., Carucci,D.J., Hoffman,S.L. and Nussenzweig,V.

TITLE Exploring the transcriptome of the malaria sporozoite stage

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9895-9900 (2001)

MEDLINE 21396555

COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org
Request for clones, please contact: Stefan Kappe,
kappe@tigr.org, med.nyu.edu Michael Heidelberger Division,
Department of Pathology New York University School of Medicine.

FEATURES

source

1..583

/organism="Plasmodium yoelii"

/strain="17XNL"

/db_xref="taxon:5861"

/clone="PYCDV83"

/clone_lib="Plasmodium yoelii sporozoite cDNA"

/dev_stage="sporozoites from salivary gland"

/lab_host="E. coli TOP10"

/note="Vector: PCR4; TA cloning; Plasmodium yoelii sporozoite cDNA library from salivary gland sporozoites 14 days post-infection"

BASE COUNT 220 a 73 c 59 g 231 t

ORIGIN

Query Match 10.1%; Score 45.2; DB 12; Length 583;

Best Local Similarity 52.1%; Pred. No. 0.6; Mismatches 93; Indels 0; Gaps 0;

Matches 101; Conservative 0;

Db 467 ATGCAGACATTTTT 480

RESULT 12

LOCUS A2541515

DEFINITION ENF0607F Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.

ACCESSION A2541515

VERSION A2541515.1 GI:11149553

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica.

REFERENCE Eukaryota; Entamoebidae; Entamoeba. 1 (bases 1 to 874)

AUTHORS Loftus,B., Van Aken,S. and Fraser,C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HMI:IMSS sheared DNA library

COMMENT Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjl@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library

Seq primer: M13-Forward

Class: Shotgun

High quality sequence start: 24

High quality sequence stop: 842.

FEATURES

source

1..874

/organism="Entamoeba histolytica"

/strain="HMI:IMSS"

/db_xref="taxon:5759"

/clone_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: pGEM1; site1; Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

BASE COUNT 275 a 84 c 89 g 426 t

ORIGIN

Query Match 10.0%; Score 44.6; DB 17; Length 874;

Best Local Similarity 51.2%; Pred. No. 0.91; Mismatches 99; Indels 0; Gaps 0;

Matches 104; Conservative 0;

245 TCTCTAAATGAGTACCTACATTTTAAATGACCTGAATGTTAGATAGTACTATATGCC 304

Db 35 TCTCTCTATGCTATATTTCTTTAAATTAATATATCTCTCAATTAATGTTT 94

305 GCTACAAAAGTAACCTTTATATTTTATATACCTTACCTCAGCAGCTATGATATA 364

Db 95 GTTTGAAATTAATTAATTAATTTGTTTCTTATTTATTCCTGATATGATGAT 154

365 AATTAACATTTTACACATATACATTAATTAATTTTATTTTCTATATGCGCAGTT 424

Db 155 ATAGAAACATGTCATTAATTTATATATTAATTAATTTCTTGACATATCTATGTTGGT 214

425 CTTTCCCTGCTTAATGAAGCT 447

Db 215 ATTATTTTATTAATGATATGAT 237

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: FR

Class: sheared ends.

FEATURES Location/Qualifiers

source

1..829

/organism="Brassica oleracea"

/strain="TO100DH3"

/db_xref="taxon:3712"

/clone="BOHJ748"

/clone_1ib="BOHJ"

/note="Vector: PHOS1, Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 315 a 86 c 102 g 326 t

ORIGIN

Query Match 9.9%; Score 44.2; DB 17; Length 829;

Best Local Similarity 58.9%; Pred. No. 1.1;

Matches 76; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```

QY 281 AATGTTAGATAGTTACTATATGCGCCGTACAAAAGGTAAAGCTTTTATATTTTATACAT 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 AATTTTGAAGAACTATATAATATGTACATCCCAATTAATATATATATATATCTTATTA 692

QY 341 TAACTTCAGCCAGCTATTTGATATAATTAACAATTTTCACACAATACATAGTTAACTA 400
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 TATTTTCGGTAAGTATATGAATTAAGAAATPACAAATTTAATATATATAAATAATTAACA 632

QY 401 TTTTATTTT 409
    | | | | |
Db 631 TAATAATTT 623

```

Search completed: February 24, 2003, 16:31:09

Job time : 669.074 secs

